DOCUMENT RESUME

ED 460 645 FL 026 605

AUTHOR Nakamura, Yuji
TITLE Assessment of English Speaking Ability.
PUB DATE 1996-07-00
NOTE 31p.; Published by Tokyo Keizai University.
PUB TYPE Journal Articles (080) -- Reports - Research (143)
JOURNAL CIT Journal of Humanitites and Natural Sciences; n102 p25-53 Jul 1996
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College Students; *Communicative Competence (Languages); *English (Second Language); Foreign Countries; Higher Education; *Language Proficiency; *Language Tests; *Oral Language; Participant Observation; Second Language Instruction; Second Language Learning; Testing
IDENTIFIERS Japan

ABSTRACT

The purpose of this research is to examine the detailed components of Japanese students' English speaking ability in terms of communicative competence by using an oral proficiency test based on Bachman's Communicative Language Ability model (included in an appendix). Eighty college students were tested on four tasks--speech making, visual material description, conversational response, and sociolinguistic competence--and their responses were rated by multiple evaluators on a scale. The results of this test show that it is possible to construct a valid, reliable, and practical test of speaking ability and shed light on native English language speaking raters' scoring of Japanese students' speaking ability. It also shows that a Japanese teacher will be capable of conducting the test in a classroom setting easily, quickly, effectively, and economically without any outside assistance. Twenty-one references and ten tables are included. (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 460 645

Assessment of English Speaking Ability

Yuji Nakamura

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Prof. Yuk. Nakamura

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

人文自然科学論集102号所載抜刷（1996年7月）

東京経済大学

2 BEST COPY AVAILABLE

CS 509 704

Assessment of English Speaking Ability

Yuji Nakamura

I. Theoretical Background and Rationale

Since the notion of communicative competence was publicized by Dell Hymes (1972), "On Communicative Competence," teachers who had previously acknowledged the limitations of Structuralism and the application of Transformational Generative Grammar to teaching, have been intrigued by the idea. The applied device of communicative competence called the "Communicative Approach" has come into vogue among English teachers around the world. In fact, the communication boom has been the center of language teaching and learning, along with the communicative approach in the field of language teaching and communicative competence (Canale and Swain, 1980; Littlewood, 1981; Richards and Schmidt, 1983). Thus, language teaching methodologies which before concentrated on linguistic aspects of a language have been largely replaced by a new dimension of language teaching focused more on communication.

The recent Gakushu-shido-yoryo (The Course of Study for Lower and Upper Secondary Schools in Japan) released by the Mombusho (The Ministry of Education 1989a; 1989b) states that one of the most important targets of English teaching in Japanese junior and senior high schools is to enhance students' oral communication ability (speaking and listening).

The research done by the JACET Study Group (1983; 1985; 1990) showed that most college students as well as college graduates want to improve their English communication ability, and in particular their speaking ability. Thus, the current trend in Japan,with regard to English Language Study, has been an emphasis on practical communication ability. Some textbooks (e.g. authorized English textbooks which have been used in junior high schools since April of 1993), teaching techniques, and materials for better communication have been developed. Also, some innovative teachers have even attacked the problems of poor communication ability by adopting unique methods such as memory games,

find your friend games, and interview games. These techniques are all introduced in Sano et al (1989), "Teaching Rules of Turn Taking" in Dornyei and Thurrell (1991), and "A Repeating Device" in Nakamura (1988).

Unfortunately, when we look at the testing side of speaking ability in Japan we see that tests of speaking are least developed among the four language skills, which include listening, speaking, reading and writing (Fukazawa 1989, Negishi 1990).

II. Purpose of the Research

The purpose of this research is to examine the detailed components of Japanese students' English speaking ability in terms of communicative competence by using this researcher's constructed oral proficiency test, which is theoretically based on Bachman's Communicative Language Ability (CLA) model (1990).(Table 1, Appendix I).

The framework proposed here was constructed in the following way. First, Bachman's Communicative Language Ability, which includes the four language skills, was modified by eliminating the strategic competence and psychophysiological mechanisms. Strategic competence, even though it is necessary in a real speaking situation and seems to cover all the processes of speaking, is not as practical or feasible in a test situation. This competence is actually a mental capacity that functions within one's mind to choose a particular strategy most appropriate to perform in a fixed situation. Therefore, it is impossible to assess the internal process function of competence. The psychophysiological mechanisms were also eliminated for practical reasons.

Second, Bachman's term "Language Competence" was renamed "Speaking Ability" since the present study only focuses on speaking. This speaking ability consists of three hypothesized competences: "Linguistic Competence," "Interactional Competence" and "Sociolinguistic Competence."

"Linguistic Competence" is just a modified version of Bachman's Organizational Competence and its existence was proven in the pilot test (Nakamura 1992b).

"Interactional Competence" was constructed to cover Bachman's Pragmatic Competence and also includes other interactional skills of speaking. The

existence of Interactional Competence was also demonstrated in the pilot test (Nakamura 1992b). However, in the present study, Interactional Competence does not completely match Bachman's "Pragmatic Competence" because it focuses more on the conversationally routine expressions rather than sociolinguistically situation-oriented expressions. Therefore, "Sociolinguistic Competence" was introduced with the hope of compensating for the sociolinguistic part of Bachman's definition of "Pragmatic Competence."

"Sociolinguistic Competence" superficially covers Bachman's overall Pragmatic Competence, but the name "Pragmatic Competence," which includes both illocutionary and sociolinguistic competence, was renamed "Sociolinguistic Competence" because of the practical limitations of illocutionary competence, where a listener's cooperative understanding is forcefully required to meet the speaker's intention of euphemistic/round-about expressions. In this study, sociolinguistic competence primarily concentrates on the sociolinguistically functional uses of spoken English in everyday situations.

The test was examined to determine if it meets the following standards:

1) The inter-rater reliability is acceptable.
2) The internal consistency reliability is acceptable.
3) The construct validity is acceptable.
4) The concurrent validity is acceptable.
5) The washback validity is acceptable.
6) The face validity is acceptable.
7) The test can be conducted and rated by Japanese teachers.
8) The test does not take an unduly long time to administer.

III. Methods

Eighty college students took the test (Table 2, Appendix I) consisting of four tasks : Task I- Speech Making, Task II- Visual-Material Description, Task III- Conversational Response Activities, Task IV- Sociolinguistic Competence Test (named Mini Contexts) (Appendix II -complete test booklet). The whole test was conducted in the language laboratory and all the responses from the students were recorded on tape.

Eleven raters (4 Japanese and 7 native English speakers), who have been

teaching English for at least one year, evaluated eighty audio tapes on which the students' responses had been recorded in the language laboratory. There was a training session before the actual evaluation of the 80 tapes. Japanese raters had a session individually and native speakers received the training in two groups. During the session, the raters were given 30 minute description covering the purpose of the research, the criteria for scoring, the content of the test and other details before listening to the sample tapes.

Each rater listened to five sample tapes from the pilot test. Of these samples, two were selected to be representative of the top 25% of the students, two of the lower 25%, and one of the middle level students. It was hoped that the raters would gain an overall criteria for each group (below average, average, above average, very good) through this training session. After listening to five sample tapes they discussed the demarcation between each tape. The raters then listened to a second set of sample tapes once again to compare the characteristics of each level and review the detailed description of the scoring criteria. When the training session was complete and all of the questions were dealt with, the raters moved on to evaluate 80 audio tapes. Each tape was evaluated by all 11 raters.*

The raters used the scoring sheet and scoring criteria designed by the present researcher to score the tapes. The tasks I and II (Speech Making and Visual-Material Description) were rated on a four point scale (below average, average, above average, very good) in each linguistic component such as grammar, vocabulary, pronunciation etc.. The difference between Bachman's CLA model and the present researcher's framework is that two scoring items Fluency and Content were added to the linguistic competence test. Fluency is the inevitable element in speaking skill and Content is necessary in the evaluation of speeches.

Conversational responses were rated on a similar four point scale (no answer, conversationally inappropriate, conversationally appropriate, very good). Sociolinguistic competence answers were rated on this same four point scale (Appendix II- Scoring Sheet and Scoring Criteria).

Note: In the end, the results of 10 raters (4 Japanese and 6 Native), out of the original 11, were used for statistical analyses. Thus, one rater's results were not

included in the statistical analysis. One reason for this exclusion from the final analysis is that this particular rater held the strong idiosyncratic principle for his scoring towards Japanese students' English speaking ability and was deaf to suggestions otherwise, possibly due to his long (20 years) teaching experience in Japan. Another reason for this exclusion is that the present researcher's rating criteria was not made clear enough to the rater although this was done almost completely successfully for the other raters. It is the opinion of this researcher that the rater's results should be eliminated from the statistical analyses in order to establish a common evaluation standard of native English speaking raters.

Data Analysis

1) Reliability

Each rater's raw score for all four tasks was added together to obtain a total score. Then, inter-rater reliability was measured through each rater's total score on the 80 tapes using Pearson's formula. The internal consistency was examined through Cronbach's Alpha to establish yet another measure of reliability.

2) Item Analysis

The item analysis (the investigation of item discrimination and item difficulty) was conducted in the following way. Item discrimination analysis was conducted with a t-test to examine if each item was functioning well enough to discriminate between the upper and lower 25% of the students. The .05 level of significance was set as the standard for the statistical significance of the probability. The operational benchmark is as follows:

(1) If more than five raters out of 10 agree in saying that an item is a non-significant item, then the item is considered to be a non-significant item.
(2) If more than three raters out of six native English speaking raters give ratings that claim that an item is non-significant, then the item is considered to be a non-significant item.

Item Difficulty is usually reported as a percentage of passing students on

each item. However, on a four point scale such as this, the item mean and standard deviation are more appropriate for checking difficulty. Therefore, in this item analysis study, the mean and the standard deviation of each rater will be examined for each item.

3) Correlation of the Four Tasks

Four tasks were examined from the viewpoint of the content validity. The four tasks should be mutually exclusive to be an appropriate composite of the entire proposed framework of speaking ability. Each task should also represent each proposed competence (Linguistic Competence, Interactional Competence, Sociolinguistic Competence).

4) Factor Analysis

Factor analysis was conducted to examine the construct validity. Since the present researcher's proposed framework consists of three competences (Linguistic Competence, Interactional Competence, Sociolinguistic Competence), this factor analysis will hopefully demonstrate if the proposed framework is reflected in the real evaluation situation.

5) Correlation between Four tasks (by Raters) and Independent Variables

The concurrent validity was examined by looking at the correlation between the four tasks (accomplished by the raters), the teachers' class grades, and the teachers' estimates.

6) Analysis of the Questionnaire Results of Three Native English Speaking Teachers Concerning Their Grading System

Three native English speaking teachers were asked about the elements which make up the students' grades in their English conversation classes. The question used for this questionnaire is as follows: " What is included in student's grades of your English conversation class?" The three teachers were also interviewed about the particular details of their grading system.

7) Analysis of the Questionnaire Results from the Students Who Took

the Speaking Test

The 80 students who took the speaking test were asked to answer a questionnaire concerning their impressions of the test from the viewpoint of their study habits and the improvement of their speaking ability. The question used is as follows: "Do you think this speaking test will change your study habits toward the improvement of your speaking ability?"

8) Examination of Practicality

The practicality and the usability of the test by Japanese teachers of English will be partially demonstrated by showing the similarities between Japanese and native English speaking raters concerning the inter-rater reliability. The test fatigue of the students and the length of the test will be examined through direct observation.

IV. Results and Discussion

Inter-Rater Reliability and Internal Consistency Reliability

The inter-rater reliability (Table 3, Appendix I) was considered acceptable (over .74 among ten raters). In addition, a reasonably high correlation (.74-.90 between individual native English speaking raters and Japanese raters) was observed. This fact indicates that Japanese teachers by themselves can conduct the test and score the results in a classroom setting with little or no help from native speakers (within the reliability range .74-.90). The internal consistency reliability (Table 4, Appendix I) (over .84) was observed and thus showed that 47 items were measuring each student's speaking ability consistently.

Item Analysis (Item Discrimination and Item Difficulty)

Item Analysis was conducted to obtain item discrimination and item difficulty. Item Discrimination showed that 39 out of the 47 total items were acceptable on statistical grounds. As was indicated in the operational benchmark of the item analysis, 8 items were not able to meet the presupposed standards. Even though these 8 items (Table 5, Appendix I) did not have

discriminating power, they were retained for psychological reasons. These 8 items give even lower level students a feeling of accomplishment in a test situation. In other words, most of the 47 items were successful in differentiating between the good group and the poor group. Item Difficulty (Table 6, Appendix I) indicates that there seems to be no substantial difference between the rating of Japanese and native English speaking raters in each item. Another discovery was that native English speaking raters were more severe about the level of Japanese students' fluency (item no. 5 and item no. 11 in the Appendix mean "fluency") than Japanese raters.

Task Correlation (Table 7 , Appendix I)

The results of task correlations demonstrate the strong correlations between Task I (Speech Making) and Task II (Visual-Material Description), and a tight correlation between Task III (Conversational Response Activities) and Task IV (Mini Contexts). The two pairs of tasks were playing a complementary role with each other with a notably strong relationship. This task correlation supported the construct validity in that the speaking ability consists of four tasks which are partially divisible.

Factor Analysis (Table 8 , Appendix I)

Even though we started with three proposed competences("Linguistic Competence," "Interactional Competence" and "Sociolinguistic Competence"), we were able to obtain two factors. We can call Factor 1 (Task I and Task II) "Linguistic Ability" because it has many substantial linguistic elements common to other tasks. We can call Factor 2 (Task III and Task IV) "Interactional-Sociolinguistic Ability" because it mainly includes situation-oriented, impromptu language ability.

Another difference between these two factors is that "Linguistic Ability" measures how well students are organizing or structuring logically connected sentences, where as "Interactional-Sociolinguistic Ability" assesses how many appropriate expressions students know. To put it in another way, the former ability deals with students' speaking ability from the view point of integration,

while the latter one examines students' speaking ability from a discrete point of view.

Yet another difference is that "Linguistic Ability" concentrates on one-way production ability by making a speech or giving a description whereas "Interactional-Sociolinguistic Ability" asks for two-way/reciprocal communication ability even in the tape-mediated situation.

The reason we changed the name from "competence" to "ability" is that both factors are dealing with the ability to use knowledge. Students should be able to do things in English with their knowledge, and the test is measuring how well students are able to do just that.

Linguistic Ability has two tasks, which include Speech Making and Visual-Material Description. Interactional-Sociolinguistic Ability also has two tasks, including Conversational Response Activities and Mini Context.

Speech Making has six scoring items (pronunciation, grammar, vocabulary, content, fluency, discourse (logicality). Visual-Material Description has these same six scoring items. Conversational Response Activities have 20 discrete point situations to clarify the situations. Mini Contexts have 15 discrete point contexts to describe the situations.

The analysis of this test's results supports the framework proposed by myself more than that of Bachman's CLA model. Bachman's Organizational Competence is close to Linguistic Ability, but this new factor proposal has a more detailed idea of speaking ability, such as fluency or content on a practical base, than Bachman's overall Communicative Language Ability. Furthermore, Bachman's Pragmatic Competence is similar to Interactional-Sociolinguistic Ability. However, this researcher's factor analysis has more concrete sub-components (items and situations) to cover the real/actual speaking ability than the theoretical model of Communicative Competence.

Finally, this newly adjusted framework, consisting of two components (Table 9, Appendix I) covers the detailed sub-components and real/practical activities of a speaking ability rather than offer the whole/overall outline of the concept of Communicative Competence. Factor Analysis, together with task correlations, helped to support this researcher's framework of speaking ability, based on Bachman's CLA model. Thus, the construct validity and content validity were each examined through factor analysis. From all of this, a refined

framework of speaking ability was established.

Concurrent Validity (Table 10 , Appendix I)

Concurrent validity, one type of criterion related validity, was investigated by comparing the results of the present test with students' grades in English Conversation classes and a teacher's estimate of students' speaking ability. The concurrent validity of this present test was supported not by class grades but by the teacher's estimates. The reason for this is because students' grades include non-language proficiency elements such as attendance, effort and participation, among other things.

Washback Validity

The washback validity was examined through a questionnaire for the students and the results supported the presupposition of washback validity. The question used for this questionnaire was as follows: " Do you think this speaking test will change your study habits toward the improvement of your speaking ability?" Students were to answer "yes" or "no" and give reasons for their answers. The result is that 80 % of them responded "yes" while only 20% of them responded "no." Typical reasons for a "yes" answer were as follows:

1) I realized that knowing English is completely different from using English or performing in English.
2) I realized that I could not use some easy expressions in a relevant context.
3) I realized that even though I could read some English, I could not speak English.
4) I realized which aspect of speaking ability I was especially poor at.
5) I realized that I could not be an effective speaker of English if I continued my present study habits, focusing on grammar-oriented textbook reading.

Reasons for "no" were as follows:

1) I did not understand what was happening in the test.
2) I was speaking to a machine and not to a human being.

Face Validity

The face validity was partially supported by this researcher's informal conversation with the students. They were excited about taking this type of unfamiliar, but seemingly authentic speaking test. Therefore, they were highly motivated to speak out in the test situation.

Examination of Practicality

Practicality was already examined along with reliability and validity. However, there are some other parts of practicality not included in the viewpoint of the test length and test booklet. This researcher's personal observation confirmed that there was no undue test fatigue for the students after the 35 minute test.

V. Implications

This new speaking test, based on Bachman's Communicative Language Ability model, was able to shed light on the most underdeveloped part of language skill testing,—the measurement of oral proficiency. It was successful in comprehensively measuring Japanese students' English speaking ability in terms of Communicative Competence and helped us understand what speaking ability truly is.

This paper has shown that it is possible to construct a valid, reliable and practical test of speaking ability, and shed light on native English speaking raters' scoring of Japanese students' speaking ability. It has also demonstrated that Japanese teachers will be capable of conducting the test in a classroom setting easily, quickly, effectively and economically, without any outside assistance.

It is hoped that a wide adoption of this test will have a positive backwash/washback effect on the study habits of students and on teaching, so that

English teaching in Japan will meet the students' needs and teach them the full extent of English communication ability (especially speaking ability).

For future research, the following points should be taken into consideration to allow for a more valid, reliable and practical test.

1) The test should be further examined in terms of further construct validation. (cf. Shohamy)
2) Other tests or estimates of speaking ability should be found so that criterion validity can be further checked. (cf. Shohamy)
3) The test should be administered to a much larger population of test takers, including as many college students as possible, to obtain more detailed information about Japanese students' speaking ability.
4) More examiners (especially Japanese) should be used so that more detailed information can be obtained from less experienced teachers. (cf. Douglas 1994)
5) The scoring criteria of this test should be compared to those of other tests, including teacher-made tests.

Note

This paper is based on my doctoral dissertation submitted to International Christian University 1993. Also a version of this article was presented at the 13 th ACROLT (Academic Committee for Research on Language Testing) Language Testing Symposium Kiryat Anavim, Israel 1994, at the 177th Monthly Meeting of the JACET (Japan Association of College English Teachers) Tokyo 1994, and at the 19th Annual Congress of the Applied Linguistics Association of Australia Melbourne, Australia 1994.

References

Bachman, L. F. 1990: *Fundamental Considerations in Language Testing.* Oxford: Oxford University Press.

Canale, M. and Swain, M. 1980: Theoretical bases of communicative

approaches to second language teaching and testing. *Applied Linguistics* 1,1-47.

Dornyei, Z. and Thurrell, S. 1991: Strategic competence and how to teach it. *ELT Journal*, 45,1,16-23.

Douglas, D. 1994: Quantity and quality in speaking test performance. *Language Testing* 11,125-144

Fukazawa, S. 1989, October: Supiikingu shido to hyoka (Teaching and evaluating speaking ability). *The English Teacher's Magazine* (pp.26-27). Tokyo: Taishukan.

Hymes, D. H. 1972: On communicative competence. In Pride, J. B. and Holmes, J., editors, *Sociolinguistics*. Harmondsworth: Penguin.

JACET (The Japan Association of College English Teachers). 1983: *A General Survey of English Language Teaching at Colleges and Universities in Japan—Teachers' Views*. Tokyo: JACET.

JACET (The Japan Association of College English Teachers). 1985: *A General Survey of English Language Teaching at Colleges and Universities in Japan—Students' Views*. Tokyo: JACET.

JACET (The Japan Association of College English Teachers). 1990: *A General Survey of English Language Teaching—College Graduates' Views*. Tokyo:
JACET.

Littlewood, W. 1981: *Communicative Language Teaching*. London: Cambridge University Press.

Mombusho (The Ministry of Education). 1989a: *Koto Gakko Gakushu Shido Yoryo* (Course of Study for High School). Tokyo: Mombusho (The Ministry of Education).

Mombusho (The Ministry of Education). 1989b: *Chugakko Gakushu Shido Yoryo* (Course of Study for Junior High School). Tokyo: Mombusho (The Ministry of Education).

Nakamura, Y. 1988: Speaking in English—linguistic problems of talkative students and psychological blocks of intelligent students. *The Chofu Bulletin* 22,99-107. Chofu Gakuen Women's Junior College.

Nakamura, Y. 1992a: Differences in native/non-native teachers'

BEST COPY AVAILABLE

evaluation of Japanese students' English speaking ability. *Cross Currents* 19.2.1-5.

Nakamura, Y. 1992b: Native speakers' assessment of Japanese students' English speaking ability: A pilot study. *The Chofu Bulletin* 25.91-124. Chofu Gakuen Women's Junior College.

Nakamura, Y. 1993: Measurement of Japanese college students' English speaking ability in a classroom setting. Unpublished Ph.D dissertation, International Christian University.

Negishi, M. 1990: Hanasu noryoku no hyoka (Assessment of speaking ability). *The Forum for English Language Teachers* (FELT) 6,39-42. Tokyo: The Institute for Research in Language Teaching.

Richards, J. c. and Schmidt, R. W. 1983: *Language and Communication.* New York and London: Longman.

Sano, M., Yoneyama, A. & Matsuzawa, S. 1989: *Kisonoryoku o tsukeru eigo shido ho* (How to help students develop their basic English ability). Tokyo: Taishukan.

Shohamy, E. 1988: A proposed framework for testing the oral language of second/foreign language learners. *Studies in Second Language Acquisition* 10,165-79.

Shohamy, E. 1994: The validity of direct versus semi-direct oral tests. *Language Testing* 11, 99-123.

APPENDIX I

Table 1

Comparison between the Bachman model and the Present Researcher's Framework

Bachman		Present Researcher
Communicative Language Ability	:	----------
Language Competence	:	Speaking Ability
1) Organizational	:	1) Linguistic
Competence	:	Competence
(1) Grammatical		------
Competence		
(2) Textual		------
Competence		
2) Pragmatic	:	2) Interactional
Competence		Competence
(1) Illocutionary		------
Competence		
(2) Sociolinguistic		------
Competence		
3) ----------	:	3) Sociolinguistic
		Competence
Strategic Competence	:	• • •
Assessment		
Planning		
Execution		
Psychophysiological	:	• • •
Mechanisms		
Productive		
Oral		
Visual		
Receptive		
Oral		
Visual		

N.B.

1) While there may be no exact one to one correspondence between Bachman's terms and the present researcher's terms, the nature of the test itself involves integration of the aspects of speaking ability. Consequently, "grammatical competence," "textual competence" or "illocutionary competence" are at least indirectly incorporated into certain of the items.

2) The asterisk (•) means that the category is not practical for assessing speaking ability.

Table 2

The Present Researcher's Theoretical Framework of Speaking Ability for the Present Test

Speaking Ability		
"Linguistic Competence"	"Interactional Competence"	"Sociolinguistic Competence"
Task 1 (Speech Making)	Task 3 (Conversational Response Activities)	Task 4 (Mini Contexts)
1) fluency	20 sentences or questions	15 mini contexts
2) discourse		
3) vocabulary		
4) grammar		
5) pronunciation		
6) content		
Task 2 (Visual-Material Description)		
1) fluency		
2) discourse		
3) vocabulary		
4) grammar		
5) pronunciation		
6) content		

Table 3

Results of Inter-Rater Reliability (Pearson's Correlation Coefficients)

	A	B	C	D	E	F	G	H	I	J	K
A											
B	.88										
C	.79	.83									
D	.91	.83	.76								
E	.90	.86	.84	.87							
F	.90	.88	.80	.89	.87						
G	.76	.74	.80	.77	.85	.77					
H	.85	.86	.78	.82	.91	.88	.87				
I	.84	.86	.85	.84	.91	.85	.88	.91			
J	.75	.77	.66	.74	.74	.73	.69	.80	.77		
K	.85	.86	.76	.84	.90	.83	.81	.87	.85	.72	

N.B. A-D: Japanese Raters n=80
E-K: Native English Speaking Raters
All coefficients were significant at the .001 level.

Table 4
Internal Consistency Reliability

rater	alpha
A	.95
B	.92
C	.94
D	.94
E	.93
F	.92
G	.84
H	.91
I	.90
K	.91

Table 5
Content of the remaining eight items

Item	Task	Description
13	III	Nice to meet you.
14	III	What is your name?
16	III	How are you?
22	III	How do you come to school?
26	III	It's a beautiful day, isn't it?
31	III	Would you like some ice cream for dessert?
35	IV	Asking for repetition
46	IV	Offering

Table 6
Item Difficulty

	J		NJ		ALL	
item	M	SD	M	SD	M	SD
1 (1)	2.07	.62	2.15	.51	2.12	.51
2 (2)	1.95	.65	1.89	.48	1.91	.49
3 (3)	1.92	.66	2.02	.47	1.98	.51
4 (4)	1.92	.68	2.04	.51	1.99	.54
5 (5)	2.03	.75	1.84	.56	1.92	.60
6 (6)	1.82	.66	1.95	.51	1.90	.54
7 (1)	2.03	.58	1.99	.47	2.00	.48
8 (2)	1.74	.55	1.77	.42	1.76	.43
9 (3)	1.82	.57	1.92	.39	1.88	.42
10 (4)	1.88	.64	1.96	.49	1.93	.51
11 (5)	1.92	.70	1.72	.55	1.80	.58
12 (6)	1.75	.64	1.81	.46	1.79	.48
13 (1)	3.64	.55	3.57	.49	3.60	.50
14 (2)	3.88	.34	3.74	.25	3.80	.28

15	(3)	3.16	.49	3.18	.52	3.17	.49
16	(4)	3.69	.36	3.63	.31	3.65	.28
17	(5)	3.01	.83	3.15	.75	3.09	.76
18	(6)	3.08	.68	3.06	.63	3.07	.63
19	(7)	2.84	.78	2.87	.68	2.85	.70
20	(8)	3.18	.61	3.09	.59	3.12	.57
21	(9)	2.83	.61	2.85	.57	2.84	.57
22	(10)	2.96	.62	2.97	.46	2.97	.50
23	(11)	3.01	.97	2.98	.96	2.99	.95
24	(12)	1.72	.97	1.85	1.06	1.80	1.01
25	(13)	2.09	.90	2.08	.89	2.08	.88
26	(14)	3.03	.55	3.10	.46	3.07	.47
27	(15)	2.30	.83	2.42	.81	2.37	.80
28	(16)	2.00	.87	2.02	.88	2.01	.86
29	(17)	2.53	.89	2.50	.87	2.51	.87
30	(18)	2.04	.84	2.30	.97	2.20	.90
31	(19)	2.99	.49	3.03	.48	3.01	.45
32	(20)	2.69	.78	2.66	.73	2.67	.72
33	(1)	2.69	.61	2.57	.54	2.62	.54
34	(2)	2.55	.52	2.45	.48	2.49	.47
35	(3)	2.59	.49	2.56	.43	2.57	.42
36	(4)	2.23	.50	2.40	.54	2.33	.49
37	(5)	2.20	.59	2.24	.50	2.22	.51
38	(6)	2.46	.45	2.45	.50	2.45	.45
39	(7)	2.28	.48	2.18	.37	2.22	.39
40	(8)	2.40	.40	2.14	.43	2.45	.39
41	(9)	2.26	.56	2.18	.48	2.21	.48
42	(10)	2.21	.50	2.13	.43	2.16	.43
43	(11)	2.60	.68	2.54	.63	2.56	.63
44	(12)	2.60	.44	2.38	.44	2.47	.41
45	(13)	2.28	.62	2.14	.51	2.20	.52
46	(14)	2.75	.62	2.78	.56	2.77	.57
47	(15)	2.14	.74	2.20	.67	2.17	.68

J: Japanese raters as one group
NJ: Native English speaking raters as one group
ALL: Combinaion of J and NJ as one group
• Numbers in the parentheses are item numbers in each task.

Table 7
Inter-Task Correlations of Ten Raters as One Group
(Using T-scores)

	Task1	Task2	Task3	Task4	All
Task1					
Task2	.78				
Task3	.57	.65			
Task4	.66	.68	.78		
All	.85	.88	.87	.90	

N.B. Each correlation was significant at the .001 level of significance.

Table 8
Factor Analysis(2 Factors)
(ten raters as one group)

item	task	Factor1	Factor2	Communality
5	1	.82	.16	.52
3	1	.80	.16	.56
6	1	.79	.14	.66
4	1	.79	.14	.64
2	1	.74	.10	.69
11	2	.73	.23	.65
8	2	.71	.19	.53
9	2	.71	.24	.54
1	1	.71	.17	.56
12	2	.70	.23	.44
7	2	.69	.23	.59
10	2	.62	.25	.54
37	4	.34	.34	.15
31	3	.25	.09	.10
25	3	.20	.58	.32
15	3	.04	.56	.08
45	4	.20	.53	.21
21	3	.15	.51	.18
23	3	.04	.49	.23
41	4	.27	.48	.25
19	3	.08	.47	.28
20	3	.16	.47	.15
43	4	.18	.47	.24
28	3	.34	.45	.21
17	3	.10	.45	.38
42	4	.24	.45	.18
44	4	.20	.44	.23
38	4	.18	.44	.32
29	3	.20	.44	.23
46	4	.02	.43	.18
27	3	.22	.43	.07
33	4	.33	.41	.18
34	4	.28	.40	.28
24	3	.23	.40	.24
13	3	.01	.39	.11
36	4	.13	.39	.17
47	4	.21	.38	.23
32	3	.21	.38	.22
22	3	.08	.38	.20
26	3	.22	.36	.16

30	3	.24	.35	.30
18	3	.28	.33	.26
40	4	.23	.33	.25
39	4	.30	.32	.23
35	4	.17	.29	.32
16	3	.11	.27	.19
14	3	.03	.25	.19
Eigen.		11.4	3.0	
Pct.of Var.		24.3	6.3	
Cum.Pct		24.3	30.6	

Table 9
Revised Framework of Speaking Ability with Tasks

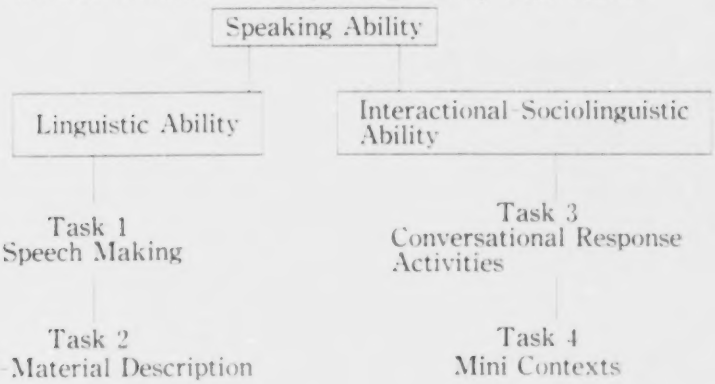


Table 10
Correlation between Teacher's Estimate and Four Tasks
(Pearson's Correlation Coefficients)

Task	TKA	TKB	TA
1	.09	.17	.71
2	.11	.10	.70
3	.17	.13	.70
4	.10	.16	.70
Total	.14	.16	.80

TKA: Conversation Class A(Grades)
TKB: Conversation Class B(Grades)
TA: Teacher's Estimate

APPENDIX II

Test Booklet, Scoring Sheet and Scoring Criteria

1). Speech (Two-minute Speech) Making Test

Directions:

1) Please choose one topic you want to talk about from among the eight topics given below.
2) Please take five minutes to prepare your speech.
3) Please give a two-minute speech about the topic you have chosen by giving the reason you chose it.

Topics:

1) My Friends 2) My Family 3) Part-time Work
4) My Hobbies 5) Traveling 6) Fashion
7) Telephone Conversations 8) College Life

2). Visual-Material Description Test

Directions:

1) Please choose one item from among the following you would like to describe.
2) Please take five minutes to prepare your description.
3) Please describe the item you chose or give as much information as possible about it within two minutes.

N.B. Some of the examples are a picture, a chart, a map, a TV program, a cartoon or an itinerary. (See Appendix III)

3). Conversational Response Test

Directions:

1) You will hear twenty questions or sentences in English each followed by a pause.
2) Please give a quick and appropriate response in English to each sentence.

Twenty recorded sentences or questions:

1) Nice to meet you.
2) What is your name?
3) Could you spell it please?
4) How are you?
5) What do you do?
6) Can you tell me the time?
7) What is the date today?
8) What is the weather like today?
9) What do you usually do on Sundays?
10) How do you come to school?
11) Thank you for everything.
12) Will you do me a favor?
13) Say hello to your family.
14) It's a beautiful day, isn't it?
15) Let's have a cup of coffee.
16) I'd like you to meet my sister.

17) I'll see you at the restaurant at six tomorrow.
18) Do you mind if I use your eraser?
19) Would you like some ice cream for dessert?
20) How about playing tennis next Sunday?

4). Sociolinguistic (Mini Contexts) Competence Test

Directions:

1) You will hear fifteen contexts in Japanese each followed by a pause.
2) Please give an appropriate response in English in each context.

Context 1 (Apologizing and making an excuse)

You are late for your class. You missed the school bus.
Please apologize and make an excuse to your teacher.

Context 2 (Complaining and requesting)

You are in a non-smoking section of a waiting room at the airport. Someone started smoking. You have a cold and a sore throat. Please complain about it and request him/her to stop it.

Context 3 (Asking for repetition)

You didn't understand what your teacher said. You want the teacher to repeat it. Please make a request to your teacher.

Context 4 (Questioning)

You want to know the train schedule. Please ask about the departure time of the next train for Kyoto at the ticket office.

Context 5 (Greeting)

You happen to meet your high school teacher (Mr. Suzuki) after a long interval. Please greet him.

Context 6 (Parting)

After talking a while, you part from your teacher. Please say "farewell" to him.

Context 7 (Disagreeing)

Your friend (Tomoko) says jogging is a healthy activity.
You don't agree with her. What do you say to her?

Context 8 (Congratulating)

Your friend's older sister won the first prize in a speech contest. Please congratulate her on her success.

Context 9 (Interrupting)

Your supervisor is working in his office. You want to interrupt him for a moment to talk with him. What do you say?

Context 10 (Warning)

Some children are playing baseball and almost break the window of your house. Please warn them.

Context 11 (Telephoning)

You are making a phone call. You want to speak to Mr. Brown. What do you say?

Context 12 (Telephoning)

You answer the phone. Someone wants to talk with your father. But he is

out now. What do you say?

Context 13 (Getting an opinion)

You want to get your friend's opinion about last week's college festival. What do you say?

Context 14 (Offering)

You want to serve something to drink to a guest at your house. Please offer something to drink.

Context 15 (Asking for information)

At a department store, please ask the receptionist where the stationery section is.

Content of Scoring Items for Speech Test and Visual Material Description Test

pronunciation
- 1) segmental features (individual sounds; vowels and consonants)
- 2) suprasegmental features (stress, rhythm, intonation)
- 3) enough volume (audible or not)
- 4) proper tone of voice

grammar
- 1) tense and aspect
- 2) noun-verb agreement
- 3) word order
- 4) noun-personal pronoun agreement

vocabulary
- 1) variety of words
- 2) word choice
- 3) idioms

content
- 1) creativity
- 2) novelty of topics

fluency
- 1) ease of speaking
- 2) speed of speaking
- 3) length, place and frequency of pauses

discourse
- 1) cohesiveness
- 2) logical combination of sentences

N.B.

All of these detailed sub-categories are employed from the results of Nakamura's research through a questionnaire given to 76 college English teachers (32 Japanese and 44 native English speaking teachers) (Nakamura, 1992).

Criteria for Four Labels in Speech Test and Visual Material Description Test

below average

Chofu students who are below average (25%)

average
Average Chofu students (25%)
above average
Chofu students who are above average (25%)
very good
Outstanding Chofu students (25%)

Criteria for Four Labels in Conversational Response Test

no answer
--there is no response
--o answer
--eaningless sounds
--incomplete phrases
--incomprehensible words/phrases
conversationally inappropriate
--the answer is conversationally inappropriate
--conversationally not acceptable
(grammatically correct or grammatically incorrect)
conversationally appropriate
--the answer is conversationally appropriate
--conversationally acceptable
(grammatically not perfct but acceptable in terms of communication)
very good
--the answer is conversationally appropriate and also very good
--grammatically correct and conversationally appropriate
--made with ease

Criteria for Four Labels in Sociolinguistic (Mini Contexts) competence Test

no answer
there is no response
sociolinguistically inappropriate
the answer is not appropriate in the assigned context
sociolinguistically appropriate
the answer is appropriate in the assigned context
very good
the answer is appropriate in the context and also very good

N.B.
The detailed explanation of each label is exactly the same as the one in the conversational response test except that "sociolinguistic" has replaced "conversational".

1. Speech Test

	below average	average	above average	very good
pronunciation				
grammar				
vocabulary				
content				
fluency				
discourse (logicality)				

2. Visual-Material Description Test

	below average	average	abore average	vevy good
pronunciation				
grammar				
vocabulary				
content				
fluency				
discourse (logicality)				

3. Conversational Response Test

	no answer	conversationally inappropriate	conversationally appropriate	very good
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

4. Sociolinguistic (Mini Contexts) Competence Test

	no answer	conversationally inappropriate	conversationally appropriate	very good
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

BEST COPY AVAILABLE

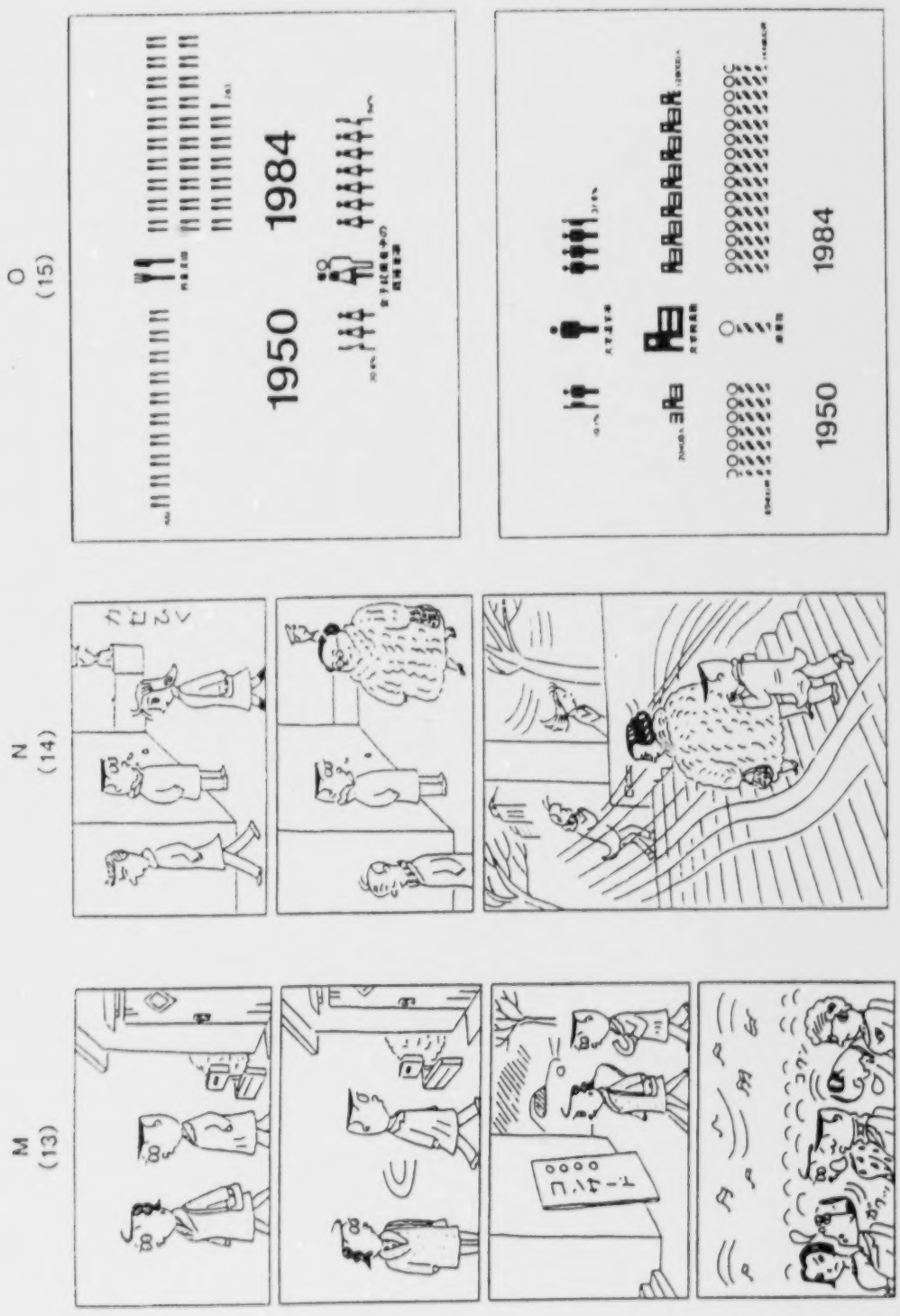
APPENDIX I

K
(11)

INTERNATIONAL MAIL			
ZONE	1	2	3
Destination / Classification	Asia (excluding U.S.S.R. in Asia)	Oceania, Middle-Near East, North America (including Alaska & Hawaii), Central America	Europe, Africa, South America
LETTERS Up to 10g	80 yen	100 yen	120 yen
For each additional 10g or fraction	60	70	100
POST CARDS	70		

L
(12)

28


M
(13)
N
(14)
O
(15)
1950
1984
1950
1984

P
(16)

JAPANESE LANGUAGE BEGINNER'S CLASS

- SMALL, FRIENDLY CLASSES
- ALL LEVELS
- PRIVATE LESSONS AVAILABLE
- FREE TRIAL LESSON
- START ANY TIME

- MORNING CLASSES 9:30—11:00
- EVENING CLASSES 19:30—21:00
- 3 TIMES A WEEK MON, WED, FRI
- LESSON FEE ¥20,000 MTH
- ENTRANCE FEE ¥20,000

RECEPTIONIST TIME MON — FRI A.M. 9:15 — P.M. 17:30

R
(18)

LOCAL TRAINS

All trains go to Cambridgeport and Milton
Change at downtown station for Rockport and Naperville

Leave airport

6 a.m. – 12 midnight every forty five minutes
12 midnight – 6 a.m. every two hours

Fare: $7.00 one way, $13.00 round trip

A.M.		P.M.	
Leave airport	Arrive downtown	Leave airport	Arrive downtown
12 midnight	12:35	12 noon	12:35
2:00	2:35	12:45	1:20
4:00	4:35	1:30	2:05
6:00	6:35	2:15	2:50
6:45	7:20	3:00	3:35
7:30	8:05	3:45	4:20

Q
(17)

Fuji (8)	Asahi (10)
6:00 Cartoon: Chibi Maruko-chan, :30 Cartoon: Sazae-san	6:00 Vival Cooking, :30 Info Variety, TV—Then & Now, :56 Weather
7:00 Cartoon: Kiteretsu Daihyakka, :30 Cartoon: Trap Ikka Monogatari, :58 News	7:00 Global Variety: Catch Me, :30 Quiz: Hint de Pinto
8:00 Info Quiz: Common Sense/Nonsense, :54 (B) News, Weather	8:00 Drama: Daihyo Tarishiman: yaku Deka, :54 News, Weather
9:00 Family Special: Kosso's Talk Show, :54 Gourmet	9:02 (B) U.S. Movie: Working Girl
10:00 Quiz: Say Quickly, :30 Shingo & Shinsuke's Talk Variety	10:54 (S) See the World by Train
11:00 (S) Music Fair, :30 FNN News, :45 Pro Baseball News	11:00 Sumo Digest, :30 Night Line ANN
12:50 Horse Racing Digest	12:00 Perle Baker Rolly, :18 Big Sports World
1:00 Theater Play Info, :25 Enterprises	1:15 (S) What's Next, :30 (S) Club Shinsuke
	2:15 (B) CNN Daywatch

S
(19)

PACKAGE A ECONO-PLAN	PACKAGE B DELUXE PLAN
Included in price: 3 days in New York City 2 nights in hotel Free admission to 3 museums	Included in price: 3 days in New York City 2 nights in hotel Free admission to 3 museums Meals (breakfast & dinner) Ballet 2 hour walking tour

30
BEST COPY AVAILABLE

T(20)

Seminar Schedule

August 1(Monday)

14:00-16:00 Registration	(Main Building)
16:00-18:00 Orientation	(Seminar Room A)
19:00-20:00 Dinner Party	(Main Dining Room)

August 2(Tuesday)

9:00-11:50 Seminar 1(English Literature)	(Seminar Room A)
12:00-13:00 Lunch	(Main Dining Room)
13:00-15:00 Seminar 2(American Literature)	(Seminar Room B)
15:30-17:30 Seminar 3(Western Culture)	(Seminar Room B)
18:00-19:00 Dinner	(Main Dining Room)
19:30-20:30 Evening Open Discussion	(Seminar Room A)

August 3(Wednesday)

9:00-11:00 Seminar 4(Japanese Culture)	(Seminar Room A)
11:30-12:30 Farewell Lunch	(Main Dining Room)

U(21)

Itinerary for Dr.Brown

July 20(Sun)	21:10	Arrives at Narita from Singapore
		Stays at Narita Tokyo International Airport Rest House.
July 21(Mon)	10:00	Mr.Yoshida meets Dr.Brown at Narita and takes him to Tokyo.
		Sightseeing in Tokyo.
	17:00	Arrives at Keio Plaza Hotel.
		Stays at Keio Plaza Hotel.
July 22(Tues)	9:00	Mr.Yoshida meets Dr.Brown and takes him to Tokyo Station.
	10:30	Leaves Tokyo for Kyoto.
		Stays in Kyoto Kokusai Hotel.
July 23(Wed)	14:30	Arrives at Haneda by ANA 306 from Kyoto. Mr.Yoshida meets him at Haneda.
	16:00	Arrives at Keio Plaza Hotel.
	18:00	Dinner with Yoshida's Family in Shinjuku.
		Stays at Keio Plaza Hotel.
July 24(Thur)	9:00	Leaves Keio Plaza Hotel for Narita.
	14:00	Leaves Narita for London by British Airways.